

5_3_Magic_BLAST_contigs_specific

July 25, 2021

Summary of magicblast

```
[1]: import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
from pandas.plotting import scatter_matrix
import pathlib
import warnings
warnings.filterwarnings('ignore')
```

```
[2]: from IPython.core.display import display, HTML
display(HTML("<style>.container { width:95% !important; }</style>"))
```

<IPython.core.display.HTML object>

```
[3]: PROJECT_CODE='PRJNA607174'
BASE_PATH = f'/mnt/1TB_0/Data/Assembly/{PROJECT_CODE}/'

dbname='nt'
kmer='k141'

#magicblast on contigs
magic_blast_sam_tail=f'_{dbname}_final_contigs_magicBLAST.sam'
f_contigs_file_tail=f'_{dbname}_magic_blast_asc_contigs.txt'
GIS,ACCESSIONS,TITLES=None,None,None
```

0.0.1 All SRA's

```
[4]: sra_list=['SRR11119759','SRR11119760','SRR11119761',\
             'SRR11119762','SRR11119763','SRR11119764',\
             'SRR11119765','SRR11119766','SRR11119767','SRR12053850']

sra_short_list=['SRR11119767','SRR12053850']
```

```
[5]: def read_gi_accession_title(gi_asc_file):
    gis=[]
    accessions=[]
```

```

titles=[]
with open(gi_asc_file, 'r') as infile:
    data = infile.readlines()
    for i in data:
        output=i.split(' ',2)
        gis.append(output[0])
        accessions.append(output[1])
        titles.append(output[2])
return gis, accessions, titles

```

```

[6]: def set_accessions():
    global GIS
    global ACCESSIONS
    global TITLES
    GIS,ACCESSIONS,TITLES=read_gi_accession_title('/mnt/1TB_ssd/Data/BLAST/nt.
    ↪gi_taxid.tsv')

```

0.0.2 Stats

```

[7]: def get_asc_descr_count(sra):
    accessions=[]
    descriptions=[]
    counts=[]
    path = BASE_PATH+sra+'/magic_blast/'
    with open(path+f'{sra}_{dbname}_{kmer}_magicBLAST_summary.txt', 'r') as f:
        lines = [line.rstrip('\n') for line in f]
        total=len(lines)-1
        for line in lines:
            if not 'database:' in line: #ignore header
                asc=line.split(None, 1)[0]
                title=line.split(None, 1)[1].split(', count:')[0]
                count=line.split(None, 1)[1].split(', count:')[1]
                accessions.append(asc)
                descriptions.append(title)
                counts.append(int(count))
    return accessions, descriptions, counts, total

```

```

[8]: def get_indexes(substr, str_list):
    index_list = []
    i = 0
    for e in str_list:
        if substr in e.lower():
            index_list.append(i)
        i +=1
    return index_list

```

```
[9]: f = open(BASE_PATH+f'/general_plots/{PROJECT_CODE}_matched_contigs.txt', "w")
for i,sra in enumerate(sra_list):
    accessions, descriptions, counts, total=get_asc_descr_count(sra)
    print(f'{sra}, matched contigs: {sum(counts)}')
    f.write(f'SRA: {sra}, matched contigs: {sum(counts)}\n')
    for d,c in zip(descriptions, counts):
        f.write(f'{d}: {c}\n')
f.close()
```

```
SRR11119759, matched contigs: 67549
SRR11119760, matched contigs: 46142
SRR11119761, matched contigs: 53040
SRR11119762, matched contigs: 67771
SRR11119763, matched contigs: 188972
SRR11119764, matched contigs: 400612
SRR11119765, matched contigs: 334800
SRR11119766, matched contigs: 378503
SRR11119767, matched contigs: 299076
SRR12053850, matched contigs: 224833
```

```
[10]: def get_desc_count(qstring, descriptions, counts, lowercase=True, ignore=None):
    qd=[]
    qc=0
    for d,c in zip(descriptions, counts):
        if lowercase:
            if ignore:
                if (qstring.lower() in d.lower()) and (ignore.lower() not in d.
↳lower()):
                    qd.append(d)
                    qc=qc+int(c)
            else:
                if qstring.lower() in d.lower():
                    qd.append(d)
                    qc=qc+int(c)
        else:
            if ignore:
                if qstring in d and ignore not in d:
                    qd.append(d)
                    qc=qc+int(c)
            else:
                if qstring in d:
                    qd.append(d)
                    qc=qc+int(c)
    #print(f'qstring: {qstring}, descriptions: {len(descriptions)}, qd: {qd},
↳counts: {len(counts)}, qc: {qc}')
    return qd, qc
```

```

[11]: def get_descr(sra):
    accessions, descriptions, counts, total=get_asc_descr_count(sra)
    #print(f'sra: {sra}, accessions: {len(accessions)}, descriptions:␣
    ↳{len(descriptions)}, counts: {len(counts)}, total: {total}')
    if total>0:
        human,humanc = get_desc_count('human', descriptions, counts)
        homo_sapiens,homo_sapiensc = get_desc_count('homo sapiens',␣
    ↳descriptions, counts)
        h_sapiens,h_sapiensc = get_desc_count('h.sapiens', descriptions, counts)
        human_contigs = human+homo_sapiens+h_sapiens
        human_counts = humanc+homo_sapiensc+h_sapiensc
        pangolin, pangolinc = get_desc_count('manis javanica', descriptions,␣
    ↳counts)
        pangolin_p, pangolin_pc = get_desc_count('manis pentadactyla',␣
    ↳descriptions, counts)
        pangolin=pangolin+pangolin_p
        pangolin_counts=pangolinc+pangolin_pc
        mouse,mousec = get_desc_count('mus musculus', descriptions, counts)
        vector,vectorc= get_desc_count('vector', descriptions, counts)
        pig,pigc = get_desc_count('sus scrofa', descriptions, counts)
        cat,catc = get_desc_count('felis catus', descriptions, counts)
        tiger,tigerc = get_desc_count('panthera tigris', descriptions, counts)
        dog,dogc = get_desc_count('canis lupus', descriptions, counts)
        virus,virusc = get_desc_count('virus', descriptions, counts,␣
    ↳ignore='retrovirus')

        mulatta,mulattac = get_desc_count('mulatta', descriptions, counts)
        troglodytes,troglodytesc = get_desc_count('troglodytes', descriptions,␣
    ↳counts)
        pongo,pongoc =get_desc_count('pongo', descriptions, counts)
        papio,papioc = get_desc_count('papio', descriptions, counts)
        mandrillus,mandrillusc =get_desc_count('mandrillus', descriptions,␣
    ↳counts)
        cercocebus,cercocebusc =get_desc_count('cercocebus', descriptions,␣
    ↳counts)
        gelada,geladac =get_desc_count('gelada', descriptions, counts)
        chlorocebus,chlorocebusc =get_desc_count('chlorocebus', descriptions,␣
    ↳counts)
        monkey =␣
    ↳mulatta+troglodytes+pongo+papio+mandrillus+cercocebus+gelada+chlorocebus
        monkey_counts =␣
    ↳mulattac+troglodytesc+pongoc+papioc+mandrillusc+cercocebusc+geladac+chlorocebusc

        mustela,mustelac= get_desc_count('mustela', descriptions, counts)

```

```

        pipistrellus, pipistrellusc = get_desc_count('pipistrellus',
↪descriptions, counts)
        rhinolophus, rhinolophusc = get_desc_count('rhinolophus', descriptions,
↪counts)
        pteropus, pteropusc = get_desc_count('pteropus', descriptions, counts)
        myotis, myotisc = get_desc_count('myotis', descriptions, counts)
        bat = pipistrellus + rhinolophus + pteropus + myotis
        bat_counts = pipistrellusc + rhinolophusc + pteropusc + myotisc

        mycoplasma, mycoplasmac = get_desc_count('mycoplasma', descriptions,
↪counts)

        lst = ['human', 'monkey', 'pangolin', 'mouse',
               'pig', 'cat', 'tiger', 'dog', 'bat', 'virus',
↪'vector', 'mycoplasma', 'mustela']
        lengths = [human_counts, monkey_counts, pangolin_counts, mousec,
                   pigc, catc, tigersc, dogc, bat_counts, virusc, vectorc,
↪mycoplasmac, mustelac]
        final_contigs = BASE_PATH + sra + '/megahit_default/final.contigs.fa'
        with open(final_contigs) as final_contigs_file:
            total_contigs = sum(1 for _ in final_contigs_file)
        res = [int(i) for i in counts]
        total_contigs_matched = sum(res)
        sra_l = [sra] * len(lst)
        fractions_matched = [human_counts/total_contigs_matched, monkey_counts/
↪total_contigs_matched, pangolin_counts/total_contigs_matched, mousec/
↪total_contigs_matched,
                           pigc/total_contigs_matched, catc/total_contigs_matched, tigersc/
↪total_contigs_matched, dogc/total_contigs_matched, bat_counts/
↪total_contigs_matched,
                           virusc/total_contigs_matched, vectorc/
↪total_contigs_matched, mycoplasmac/total_contigs_matched, mustelac/
↪total_contigs_matched]
        pct_matched = [round(i * 100, 2) for i in fractions_matched]
        df = pd.DataFrame(list(zip(sra_l, lst, lengths, pct_matched)),
                           columns=['SRA', 'Name', 'count', 'pct_matched'])

        return df
    return None

```

```

[12]: sns.set(rc={"figure.figsize": (4, 4)})
def plot_df(df, sra):
    ax = df.plot(x='Name', y='count', kind='bar')
    ax.set_title(sra, fontsize=12)
    #ax.set_yscale('log')
    #ax.set_ylim([0, df['count'].max()+10])
    ax.set_ylim(bottom=0)

```

```

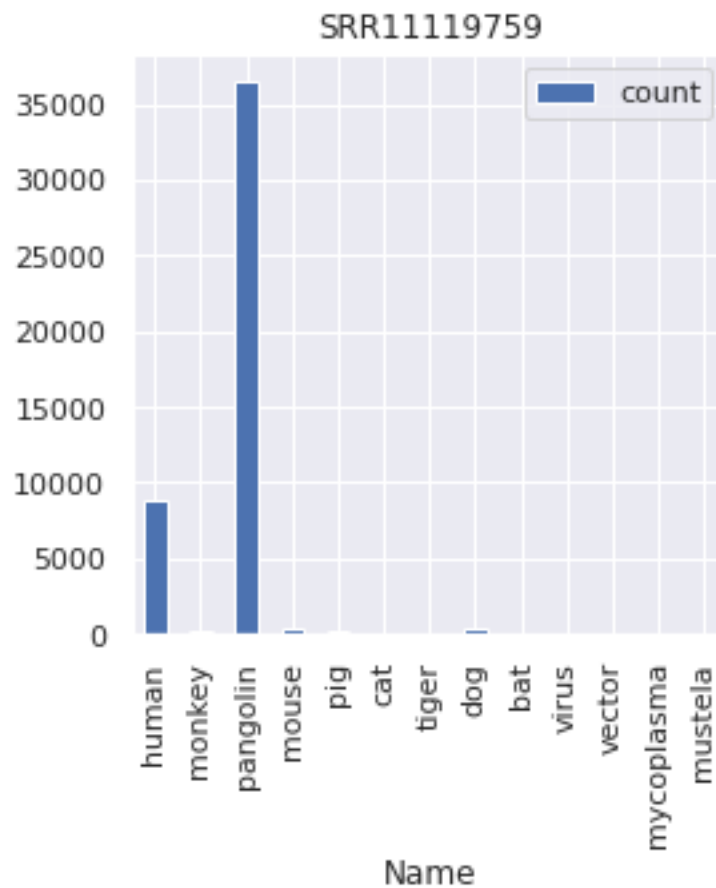
fig = plt.gcf()
fig.savefig(BASE_PATH+sra+'/magic_blast/
↳'+f'{sra}_{dbname}_{kmer}_magicBLAST_species.png', bbox_inches="tight")

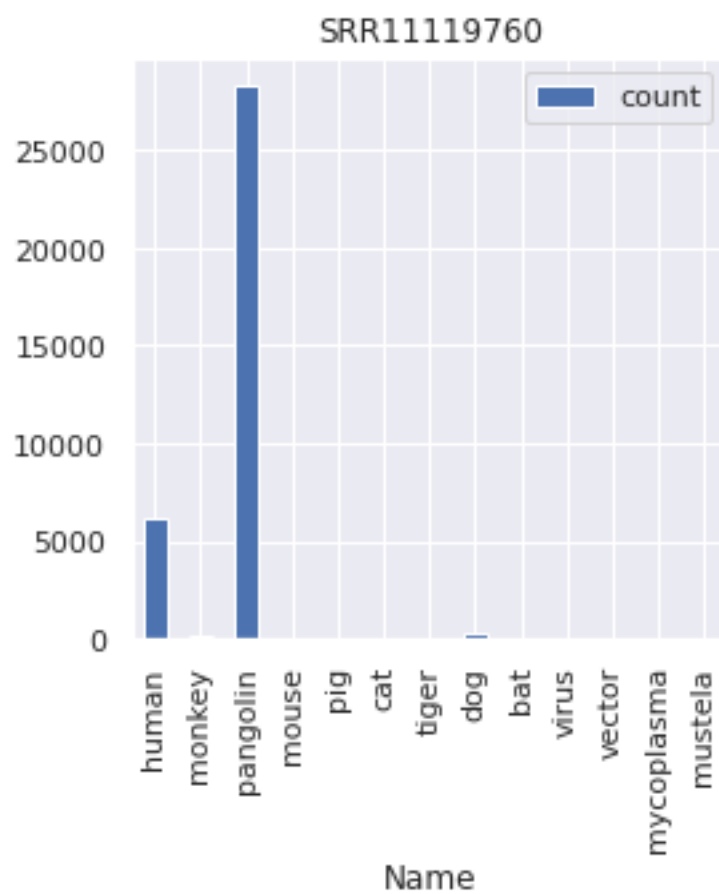
```

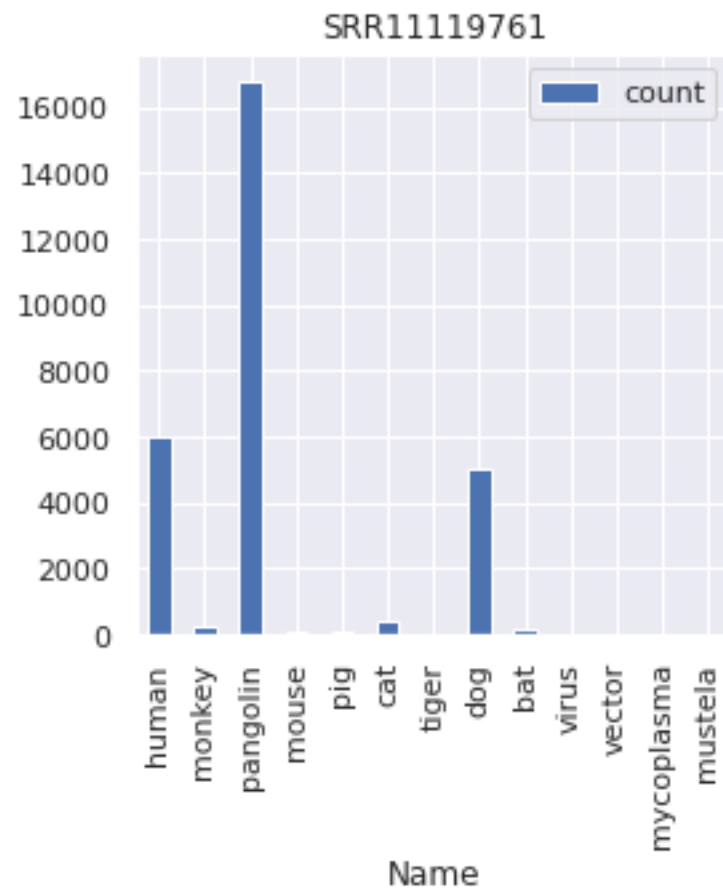
```

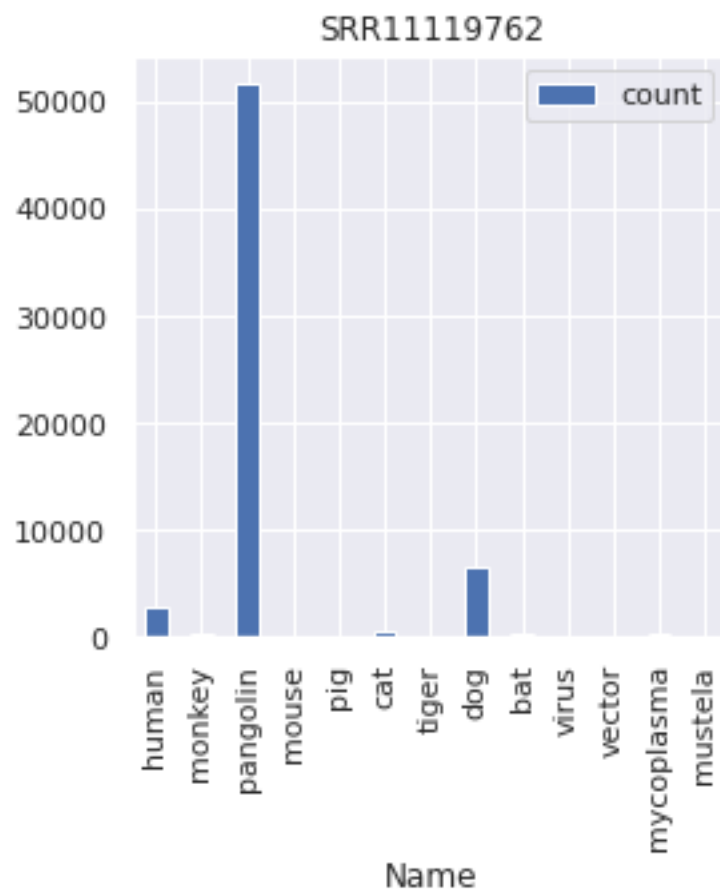
[13]: for sra in sra_list:
    try:
        df=get_descr(sra)
        df.to_csv(BASE_PATH+sra+'/magic_blast/
↳'+f'{sra}_{dbname}_{kmer}_magicBLAST_species_df.csv')
        plot_df(df, sra)
    except FileNotFoundError:
        pass
    except AttributeError:
        pass

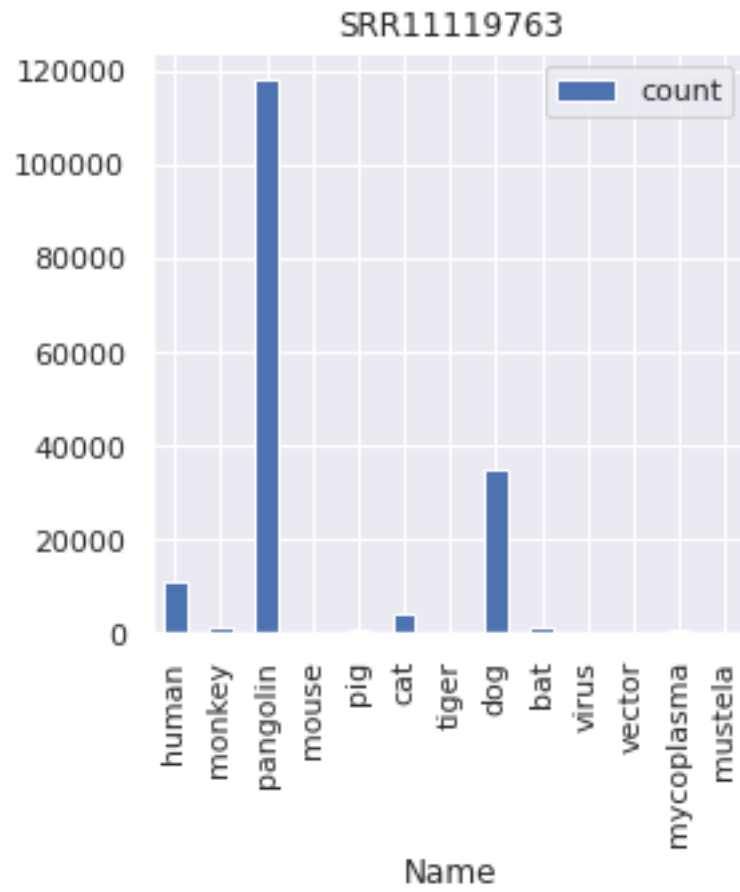
```

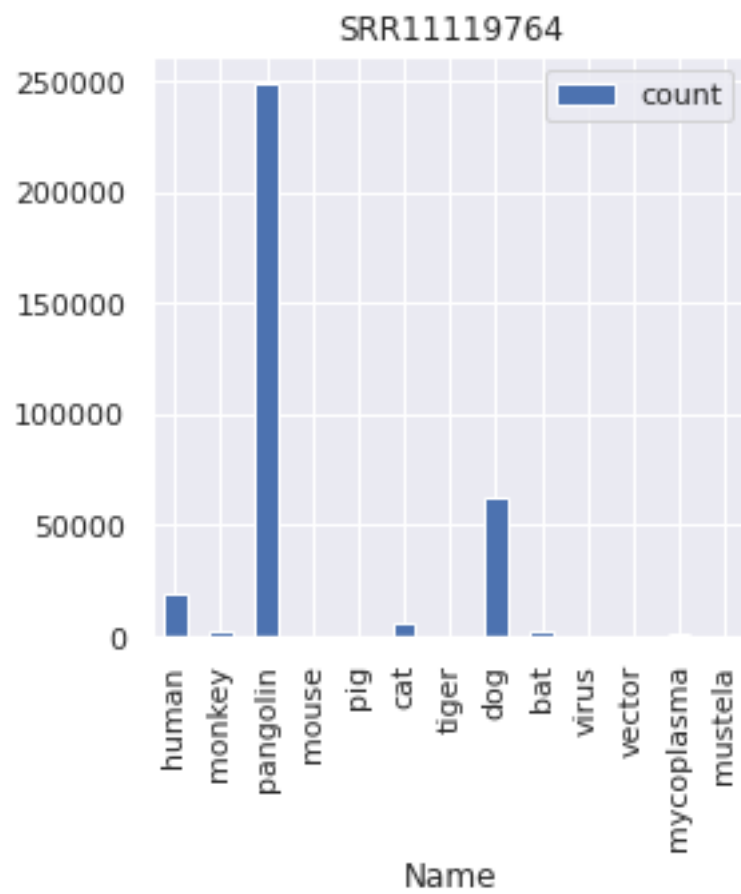


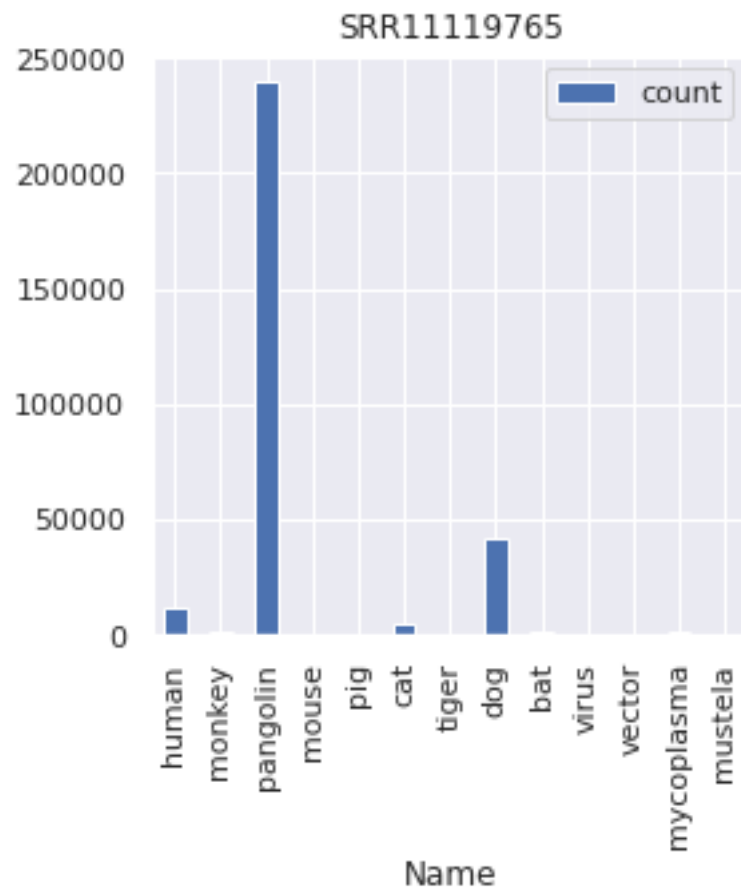


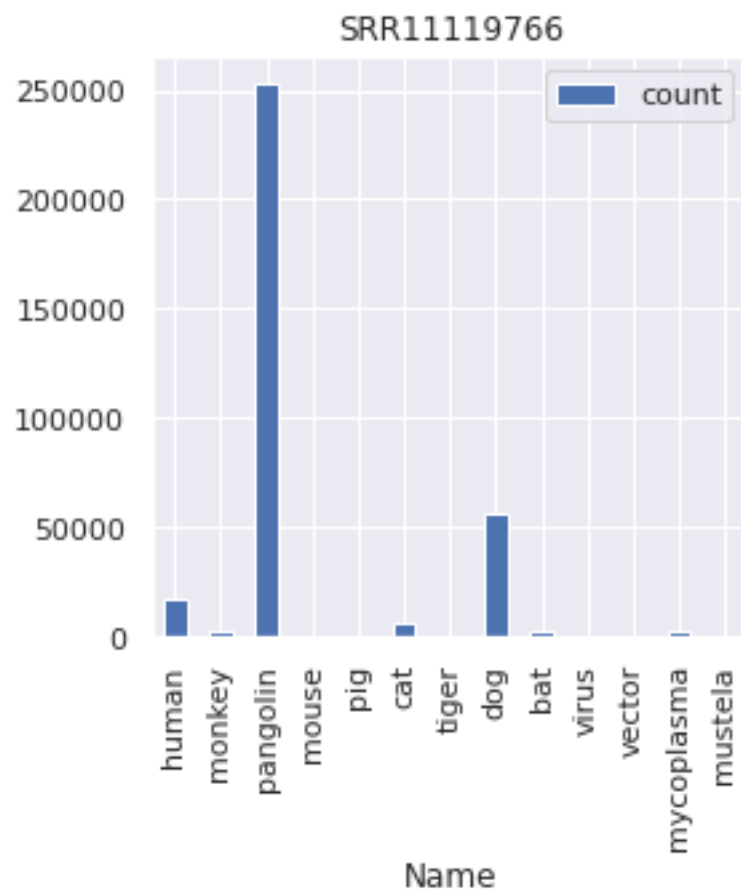


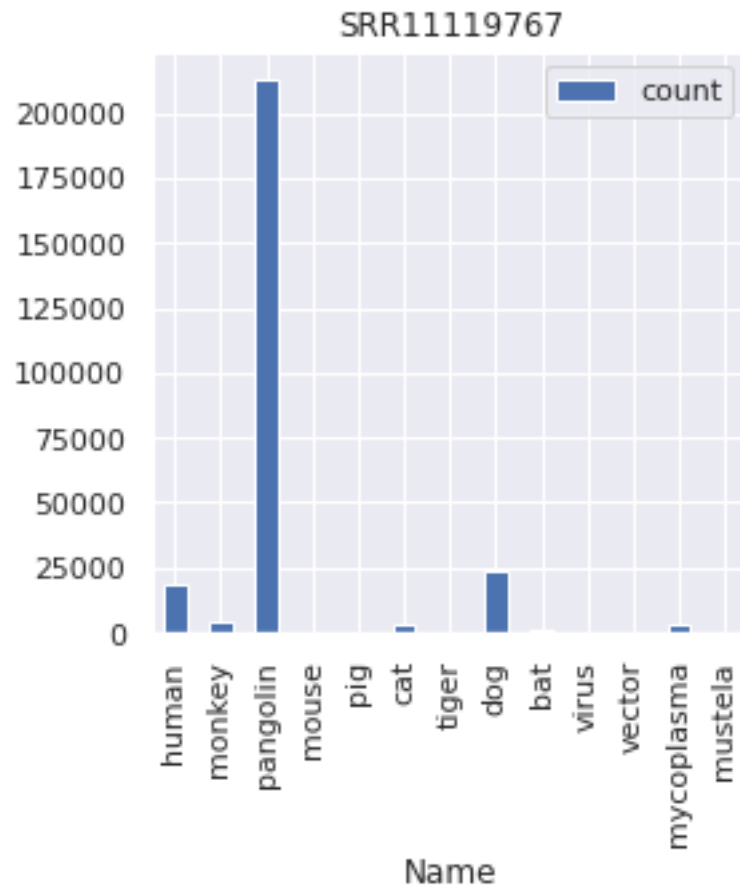


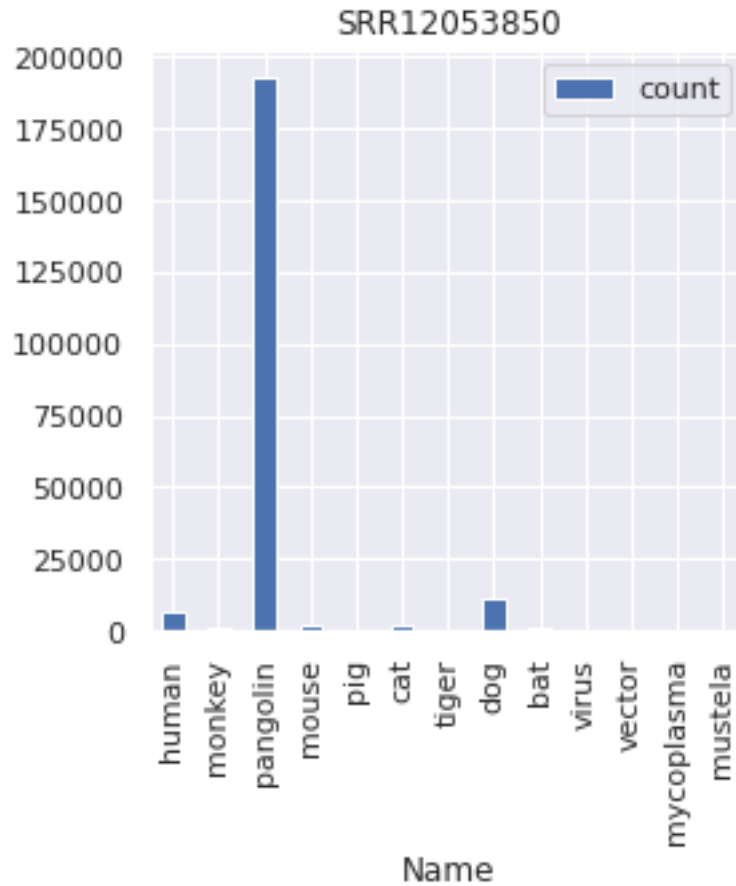












```
[14]: pathlib.Path(BASE_PATH+'general_plots/').mkdir(exist_ok=True)
```

```
def multi_plot():
    fig, axis = plt.subplots(2, 5, figsize=(16,8))
    fig.suptitle('Contig counts')
    sra=sra_list[0]
    df=get_descr(sra)
    axis[0,0].bar(df['Name'],df['count'])
    axis[0,0].tick_params(axis='x', rotation=90)
    axis[0,0].set_ylabel('count')
    axis[0,0].title.set_text(sra)

    sra=sra_list[1]
    df=get_descr(sra)
    axis[0,1].bar(df['Name'],df['count'])
    axis[0,1].set_ylabel('count')
    axis[0,1].tick_params(axis='x', rotation=90)
    axis[0,1].set_ylabel('count')
    axis[0,1].title.set_text(sra)
```

```

sra=sra_list[2]
df=get_descr(sra)
axis[0,2].bar(df['Name'],df['count'])
axis[0,2].set_ylabel('count')
axis[0,2].tick_params(axis='x', rotation=90)
axis[0,2].set_ylabel('count')
axis[0,2].title.set_text(sra)

sra=sra_list[3]
df=get_descr(sra)
axis[0,3].bar(df['Name'],df['count'])
axis[0,3].set_ylabel('count')
axis[0,3].tick_params(axis='x', rotation=90)
axis[0,3].set_ylabel('count')
axis[0,3].title.set_text(sra)

sra=sra_list[4]
df=get_descr(sra)
axis[0,4].bar(df['Name'],df['count'])
axis[0,4].set_ylabel('count')
axis[0,4].tick_params(axis='x', rotation=90)
axis[0,4].set_ylabel('count')
axis[0,4].title.set_text(sra)

sra=sra_list[5]
df=get_descr(sra)
axis[1,0].bar(df['Name'],df['count'])
axis[1,0].set_ylabel('count')
axis[1,0].tick_params(axis='x', rotation=90)
axis[1,0].set_ylabel('count')
axis[1,0].title.set_text(sra)

sra=sra_list[6]
df=get_descr(sra)
axis[1,1].bar(df['Name'],df['count'])
axis[1,1].set_ylabel('count')
axis[1,1].tick_params(axis='x', rotation=90)
axis[1,1].set_ylabel('count')
axis[1,1].title.set_text(sra)

sra=sra_list[7]
df=get_descr(sra)
axis[1,2].bar(df['Name'],df['count'])
axis[1,2].set_ylabel('count')
axis[1,2].tick_params(axis='x', rotation=90)

```



```

axis[1,2].set_ylabel('count')
axis[1,2].title.set_text(sra)

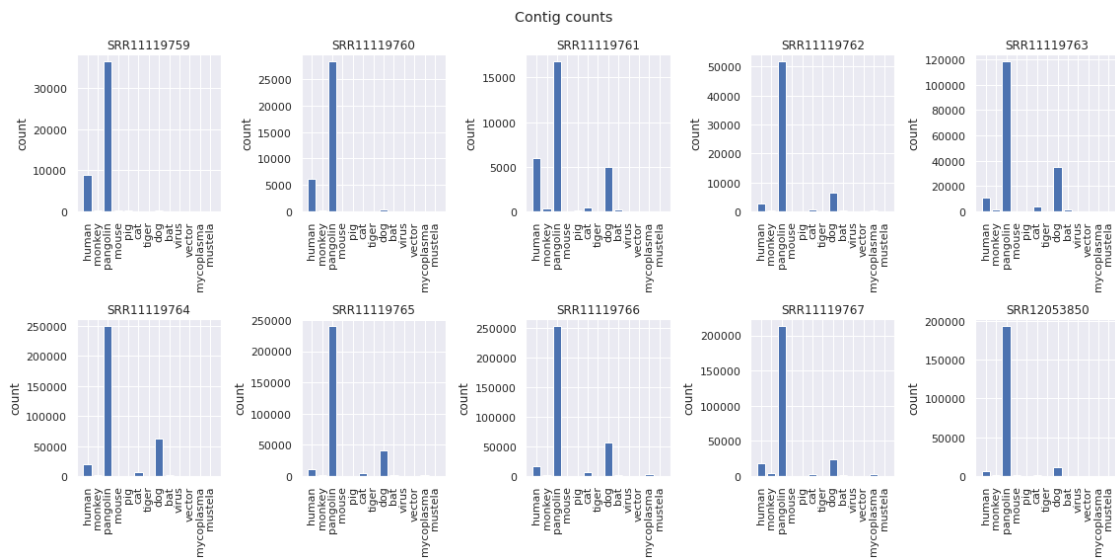
sra=sra_list[8]
df=get_descr(sra)
axis[1,3].bar(df['Name'],df['count'])
axis[1,3].set_ylabel('count')
axis[1,3].tick_params(axis='x', rotation=90)
axis[1,3].set_ylabel('count')
axis[1,3].title.set_text(sra)

sra=sra_list[9]
df=get_descr(sra)
axis[1,4].bar(df['Name'],df['count'])
axis[1,4].set_ylabel('count')
axis[1,4].tick_params(axis='x', rotation=90)
axis[1,4].set_ylabel('count')
axis[1,4].title.set_text(sra)

fig.tight_layout()
fig.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_count_matrix.png',↳
↳bbox_inches="tight")
plt.show()

```

[15]: multi_plot()



```
[16]: frames=[]
      for i,sra in enumerate(sra_list):
          try:
              df = pd.read_csv(BASE_PATH+sra+'/magic_blast/
→'+f'{sra}_{dbname}_{kmer}_magicBLAST_species_df.csv')
              x = df.Name.astype('category')
              df['species_uid'] =x.cat.codes
              df['SRA_val'] =df.SRA.str.strip('SRR')
              df['SRA_val'] = pd.to_numeric(df['SRA_val'])
              frames.append(df)
          except FileNotFoundError:
              pass
      df_sra = pd.concat(frames)
```

```
[17]: df_sra.drop(columns=['Unnamed: 0'],inplace=True)
```

```
[18]: df_sra.head()
```

```
[18]:
```

| | SRA | Name | count | pct_matched | species_uid | SRA_val |
|---|-------------|----------|-------|-------------|-------------|----------|
| 0 | SRR11119759 | human | 8837 | 13.08 | 3 | 11119759 |
| 1 | SRR11119759 | monkey | 226 | 0.33 | 4 | 11119759 |
| 2 | SRR11119759 | pangolin | 36505 | 54.04 | 8 | 11119759 |
| 3 | SRR11119759 | mouse | 343 | 0.51 | 5 | 11119759 |
| 4 | SRR11119759 | pig | 191 | 0.28 | 9 | 11119759 |

```
[19]: df_sra.Name.unique()
```

```
[19]: array(['human', 'monkey', 'pangolin', 'mouse', 'pig', 'cat', 'tiger',
          'dog', 'bat', 'virus', 'vector', 'mycoplasma', 'mustela'],
          dtype=object)
```

0.0.3 All nt database matches

```
[20]: total_dict={}
      asc_desc={}
      for sra in sra_list:
          accessions, descriptions, counts, total=get_asc_descr_count(sra)
          for asc,desc,cnt in zip(accessions,descriptions, counts):
              if asc in total_dict:
                  total_dict[asc]+=int(cnt)
              else:
                  total_dict[asc]=int(cnt)
              if asc not in asc_desc:
                  asc_desc[asc]=desc
```

```
[21]: assert len(total_dict)==len(asc_desc)
```

```
[22]: len(asc_desc)
```

```
[22]: 201789
```

Print out the 100 most common nt database classification for all contigs in the project

```
[23]: top_n=[]
def get_top_n_sp(total_dict, max_num, ignore=''):
    listofTuples = sorted(total_dict.items() , reverse=True, key=lambda x: x[1])
    for i, elem in enumerate(listofTuples):
        if i<max_num:
            if ignore=='':
                top_n.append(asc_desc[elem[0]]+ ": " + str(elem[1]) )
            else:
                if ignore.lower() not in asc_desc[elem[0]].lower():
                    top_n.append(asc_desc[elem[0]]+ ": " + str(elem[1]))
        else:
            break
    return top_n
```

```
[24]: #ignoring Pangolins
top_n =get_top_n_sp(total_dict, max_num=10, ignore='Manis')
for n in top_n:
    print(n)
```

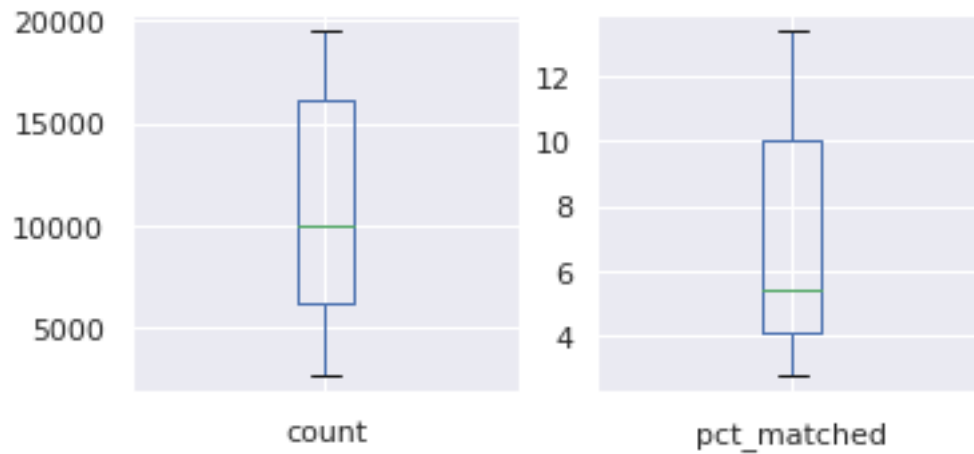
Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538

0.0.4 Human

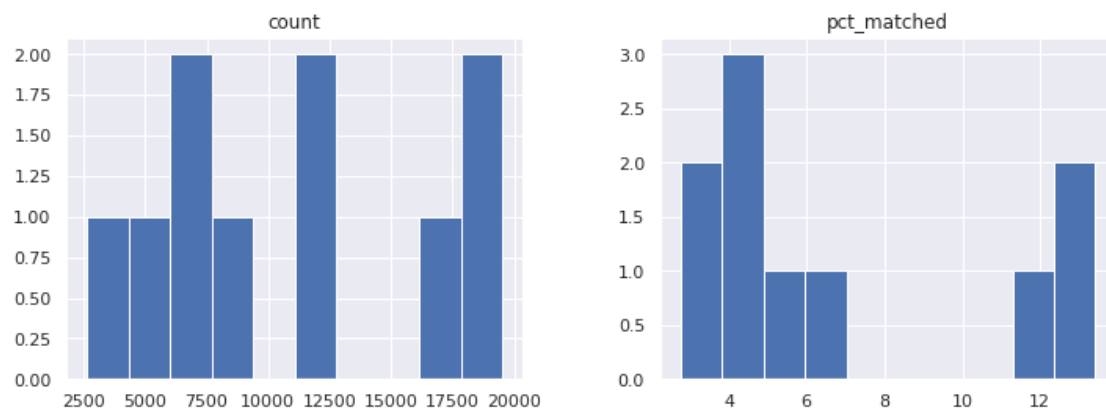
```
[25]: df=df_sra[df_sra['Name'].str.contains('human')]
```

```
[26]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
→sharey=False, figsize=(12,12))
```

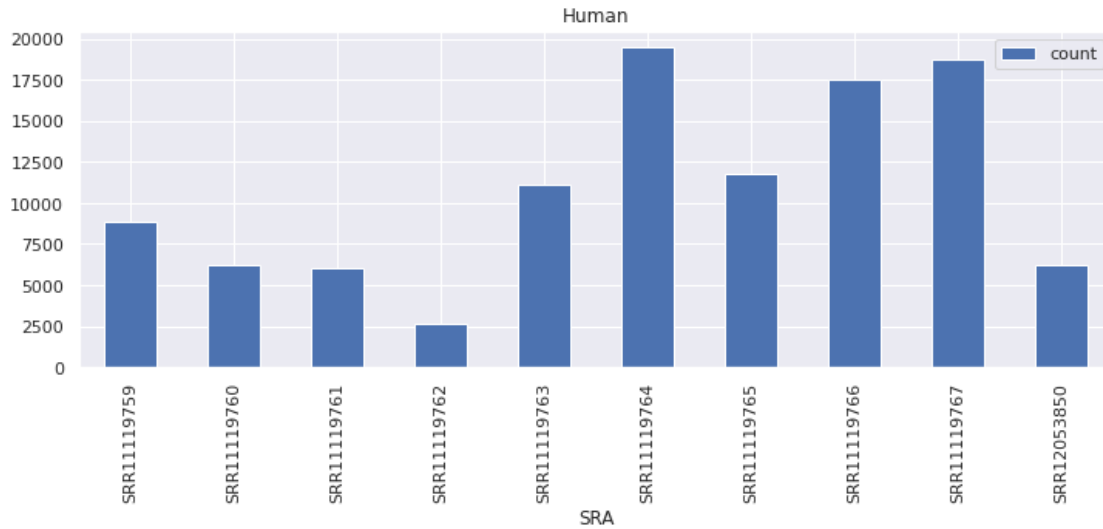
```
[26]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched      AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



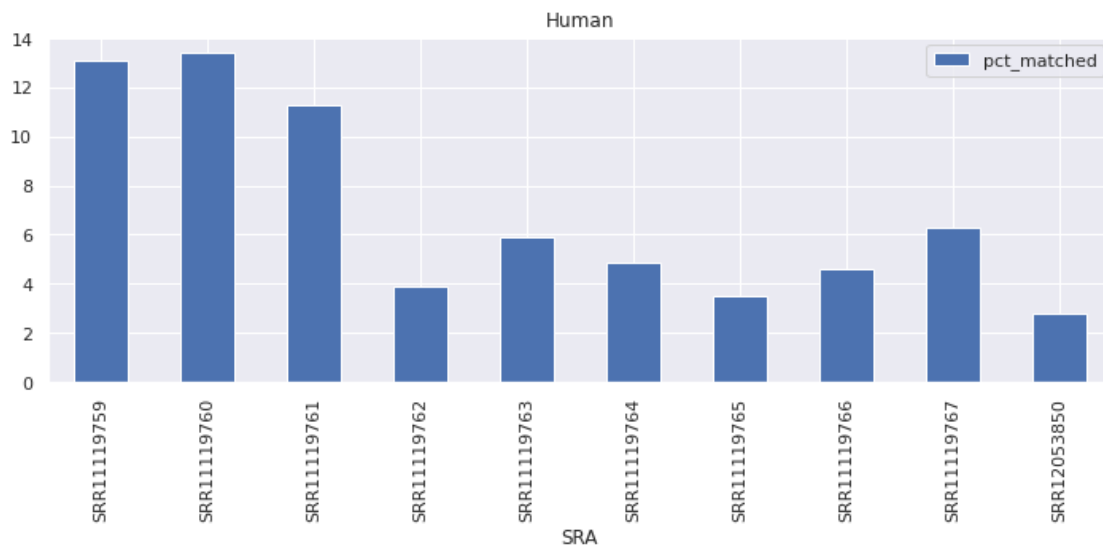
```
[27]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[28]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title('Human', fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_human.png',
↳bbox_inches="tight")
plt.show()
```



```
[29]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title('Human', fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_human_pct.png',
↳bbox_inches="tight")
plt.show()
```



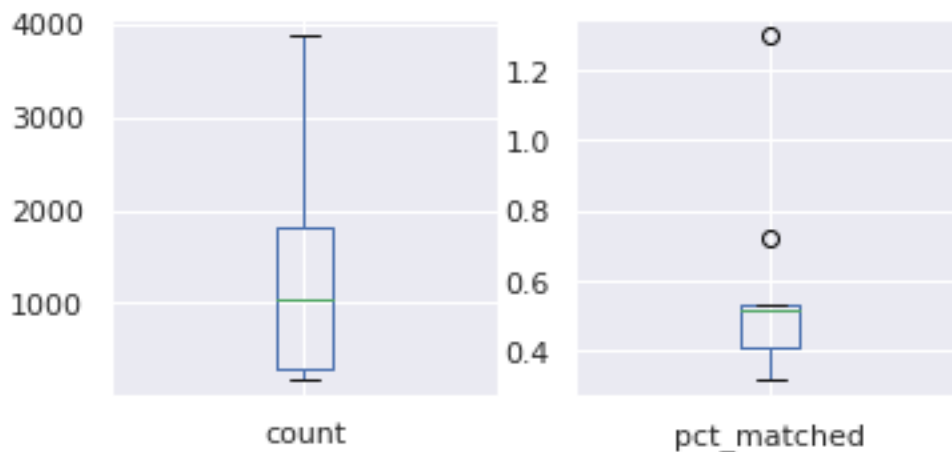
0.0.5 Monkey

```
[30]: s_name='monkey'
```

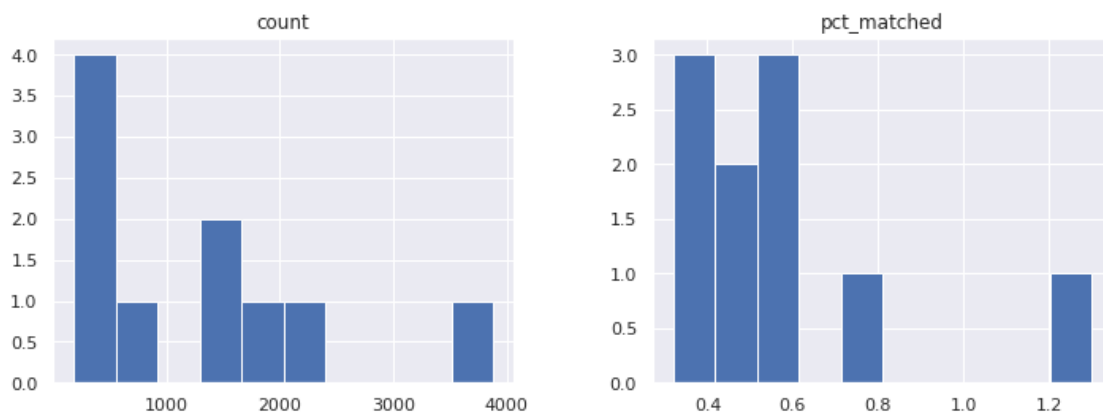
```
[31]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[32]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
↪sharey=False, figsize=(12,12))
```

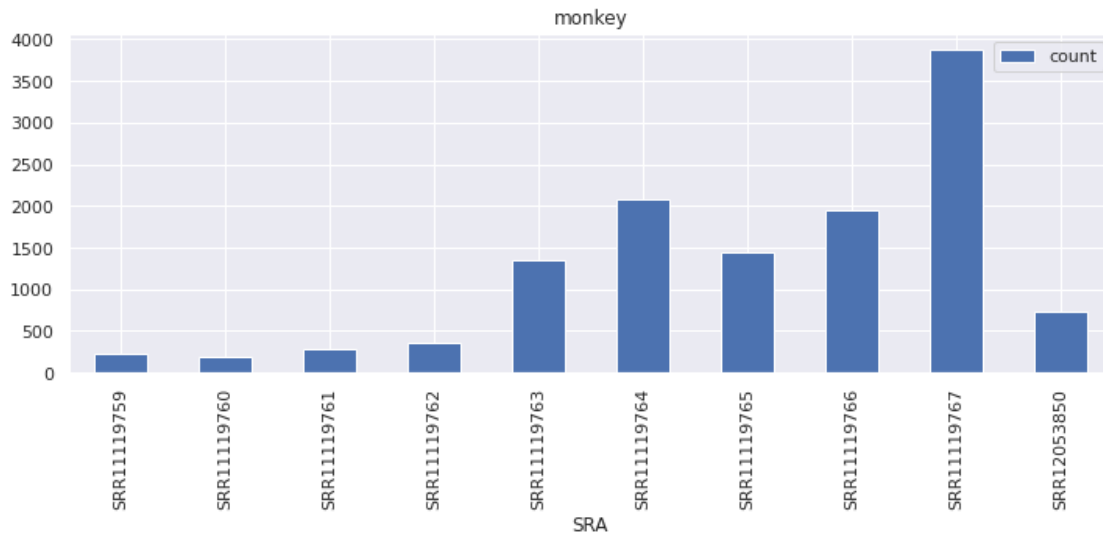
```
[32]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched      AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



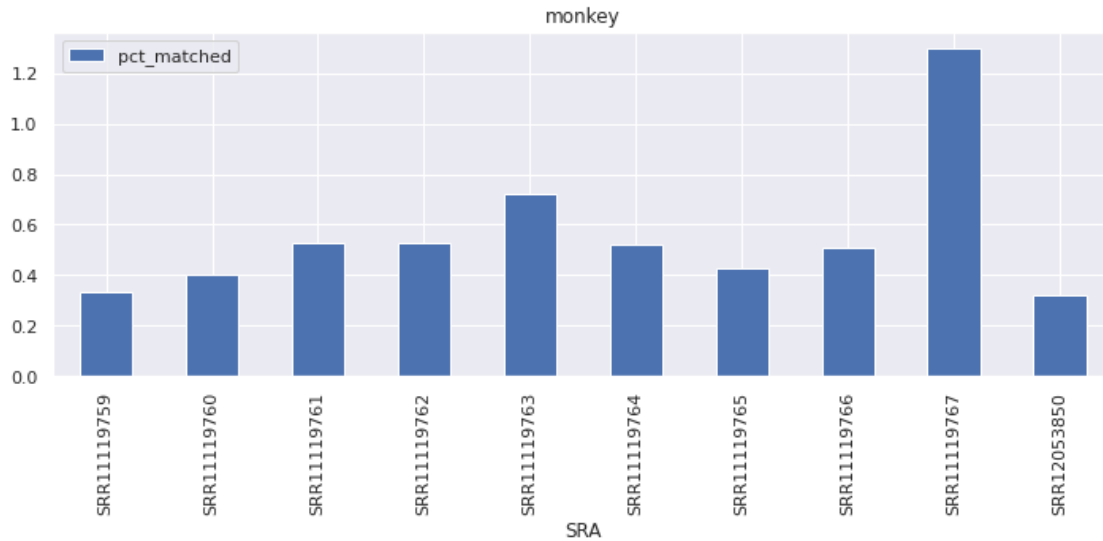
```
[33]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[34]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_monkey.png',
↳bbox_inches="tight")
plt.show()
```



```
[35]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_monkey_pct.png',
↳bbox_inches="tight")
plt.show()
```



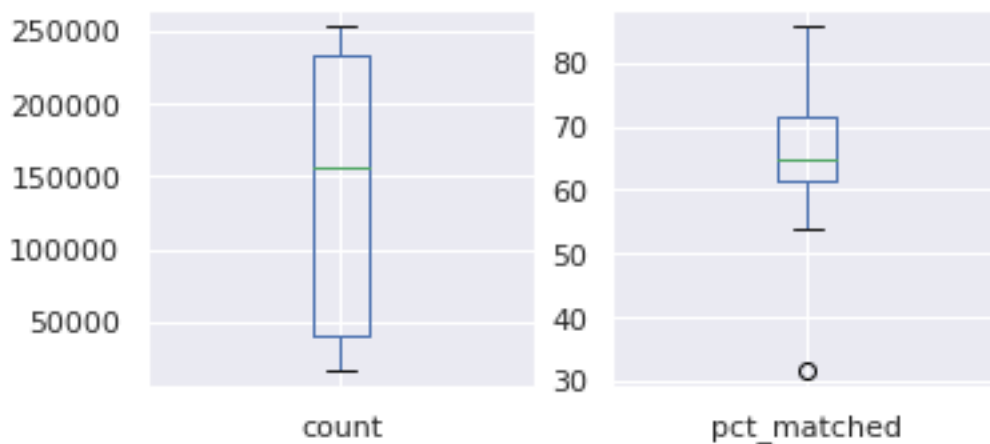
0.0.6 Pangolin

```
[36]: s_name='pangolin'
```

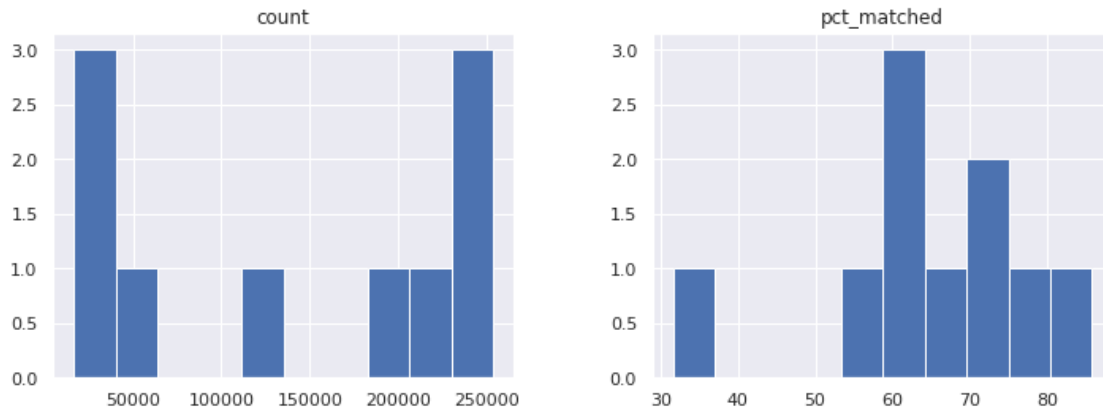
```
[37]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[38]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
↪sharey=False, figsize=(12,12))
```

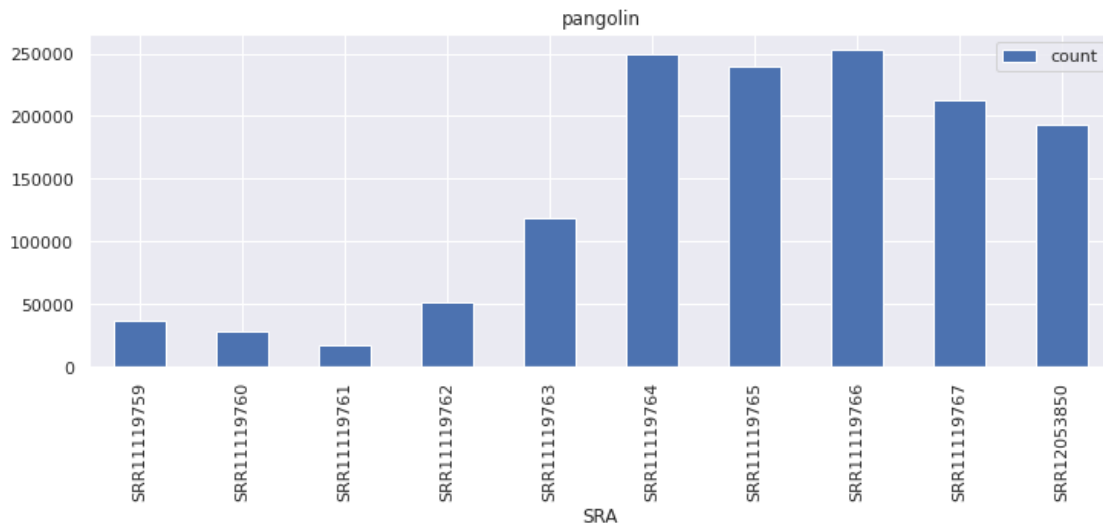
```
[38]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched        AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```




```
[39]: df_box.hist(figsize=(12,4))
plt.show()
```

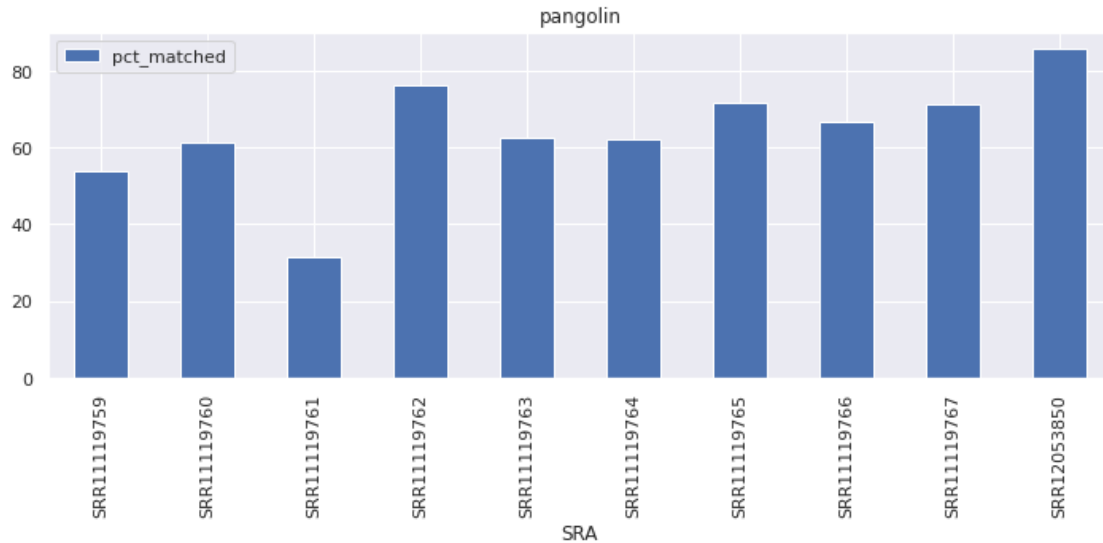


```
[40]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_pangolin.png',
↳bbox_inches="tight")
plt.show()
```



```
[41]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
```

```
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_pangolin_pct.png',
↳bbox_inches="tight")
plt.show()
```



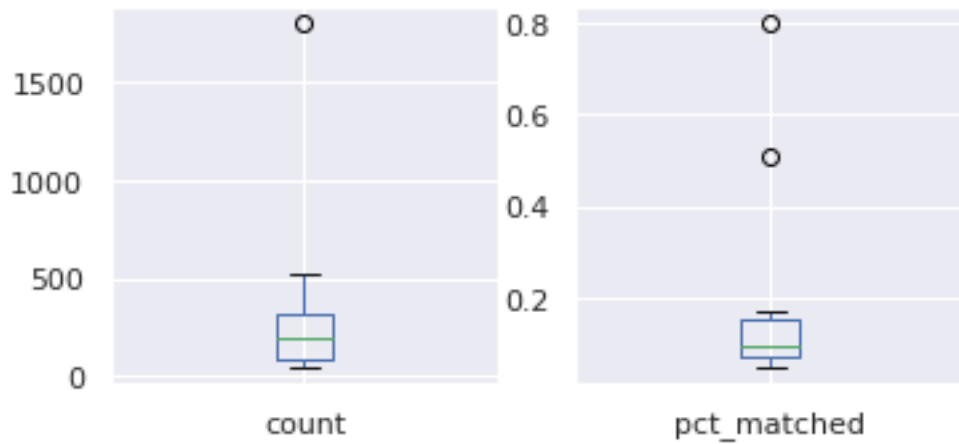
0.0.7 Mouse

```
[42]: s_name='mouse'
```

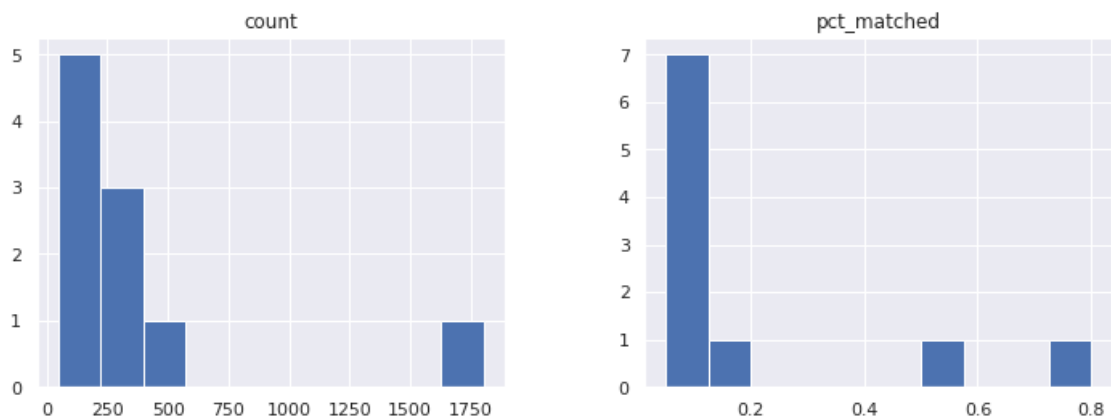
```
[43]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[44]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
↳sharey=False, figsize=(12,12))
```

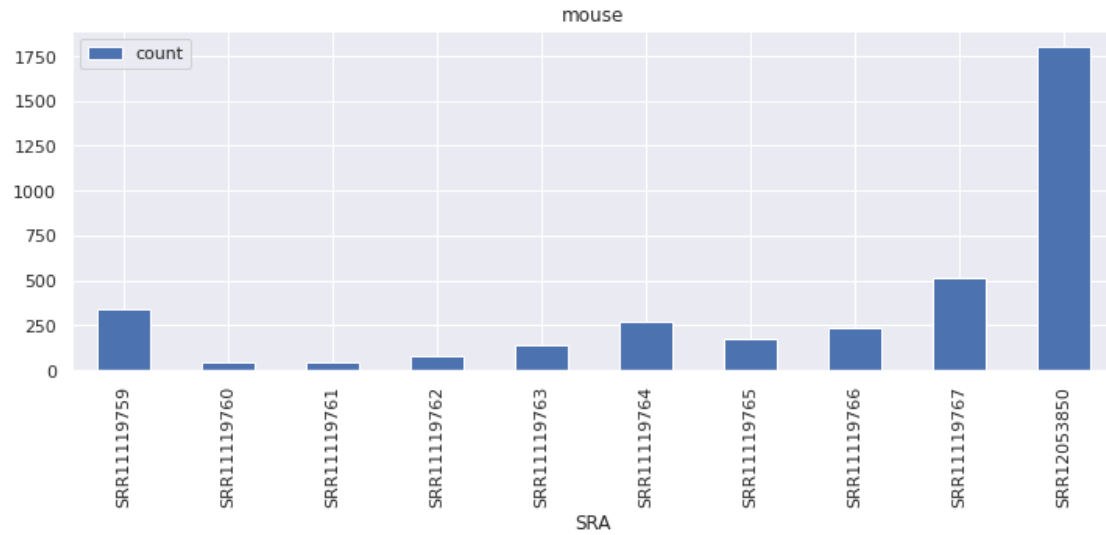
```
[44]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched        AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



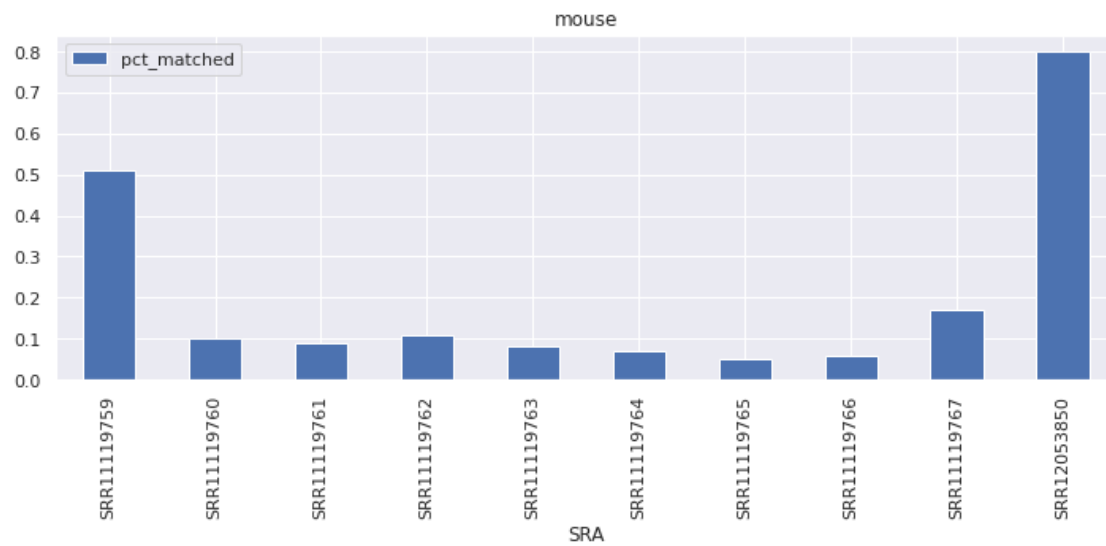
```
[45]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[46]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_mouse.png',
↳bbox_inches="tight")
plt.show()
```



```
[47]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_mouse_pct.png',
↳bbox_inches="tight")
plt.show()
```



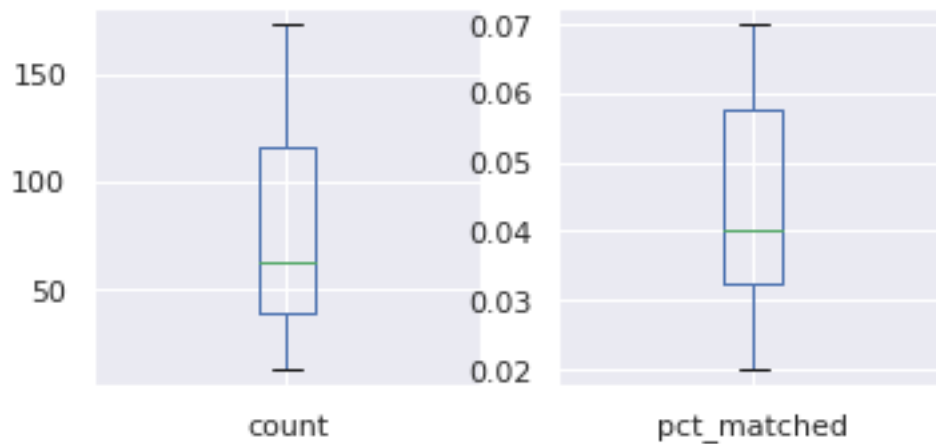
0.0.8 virus

```
[48]: s_name='virus'
```

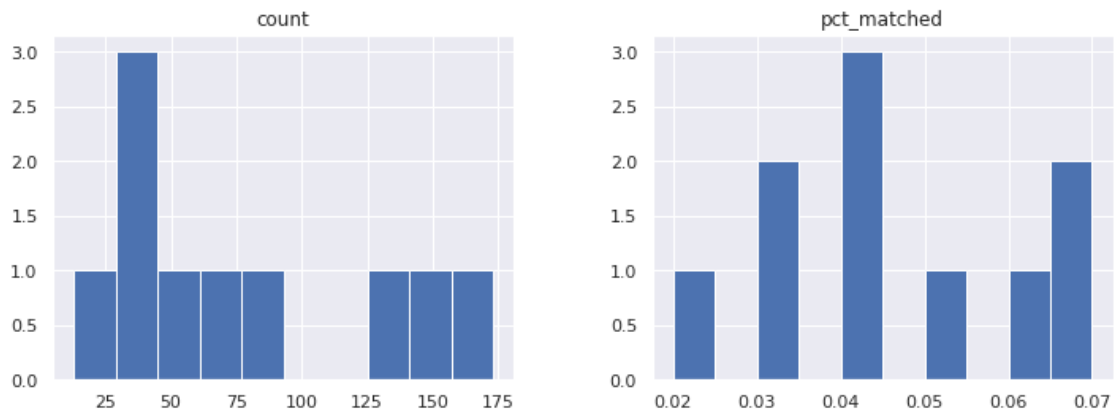
```
[49]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[50]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
→sharey=False, figsize=(12,12))
```

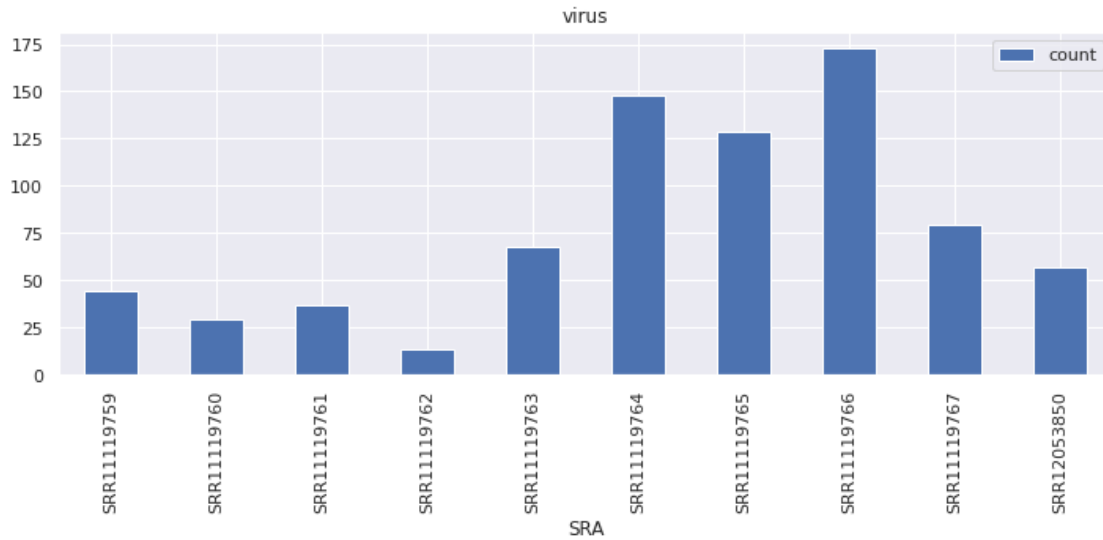
```
[50]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched      AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



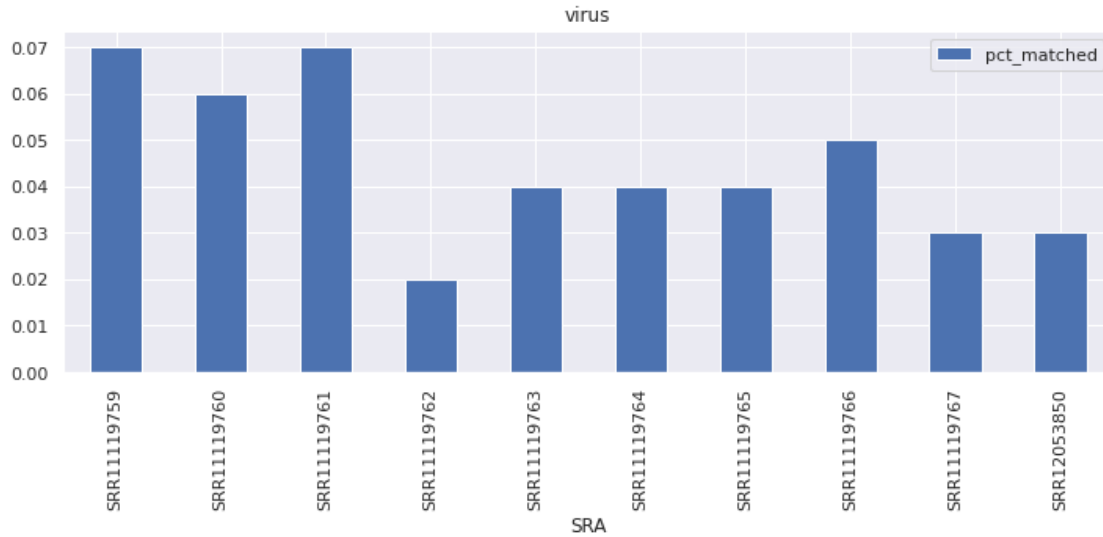
```
[51]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[52]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_virus.png',
↳bbox_inches="tight")
plt.show()
```



```
[53]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_virus_pct.png',
↳bbox_inches="tight")
plt.show()
```



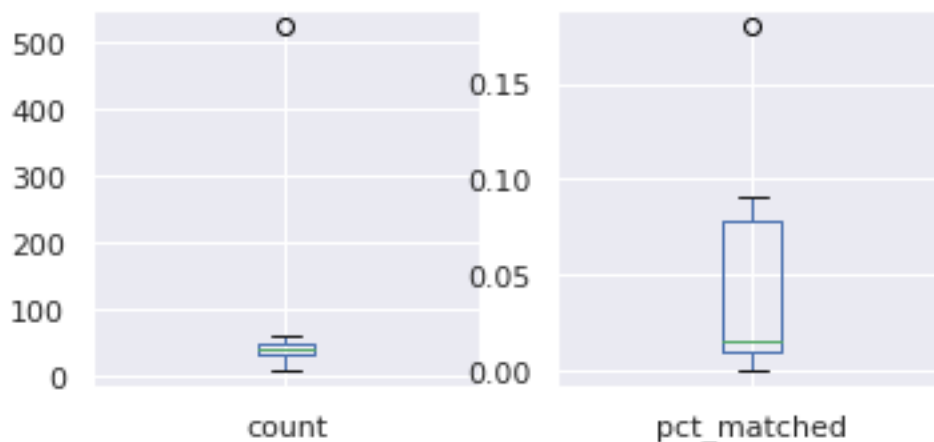
0.0.9 vector

```
[54]: s_name='vector'
```

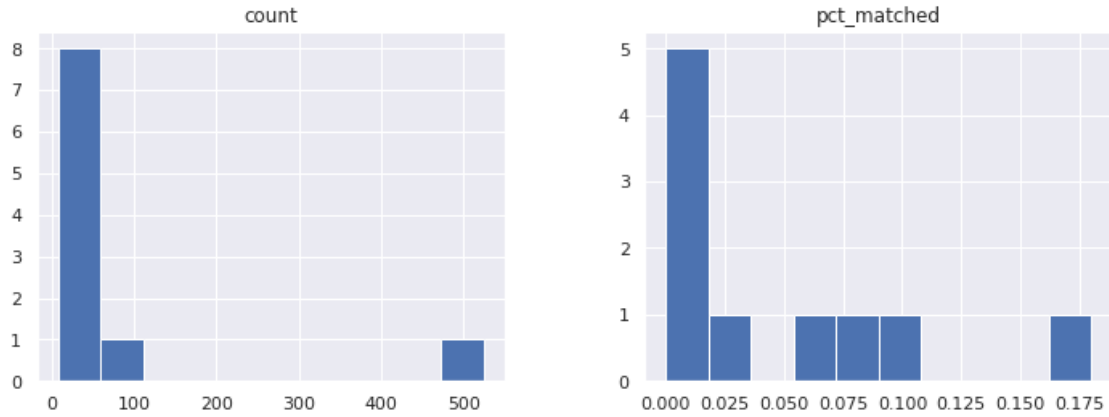
```
[55]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[56]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
↪sharey=False, figsize=(12,12))
```

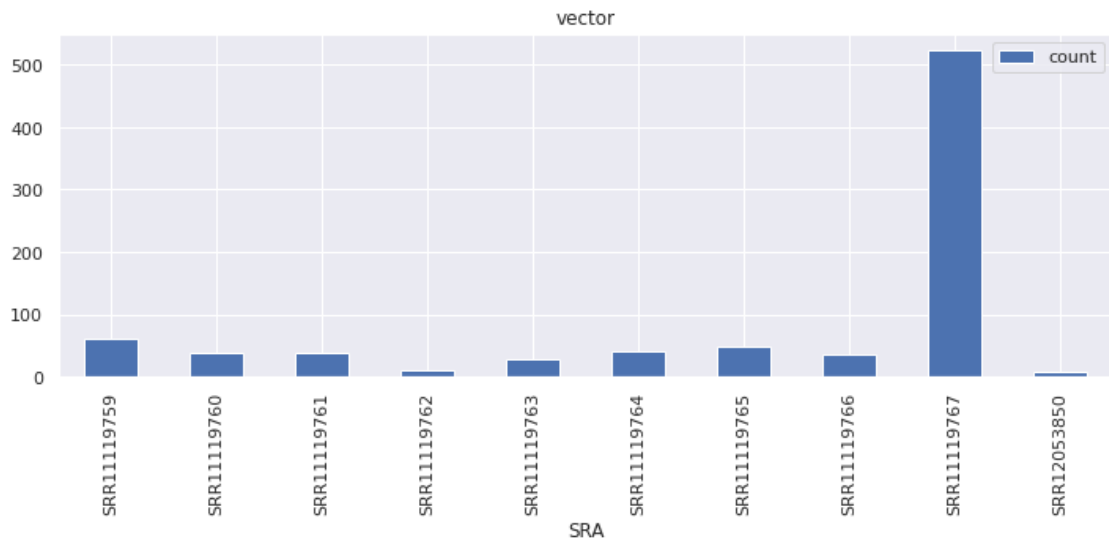
```
[56]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched        AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



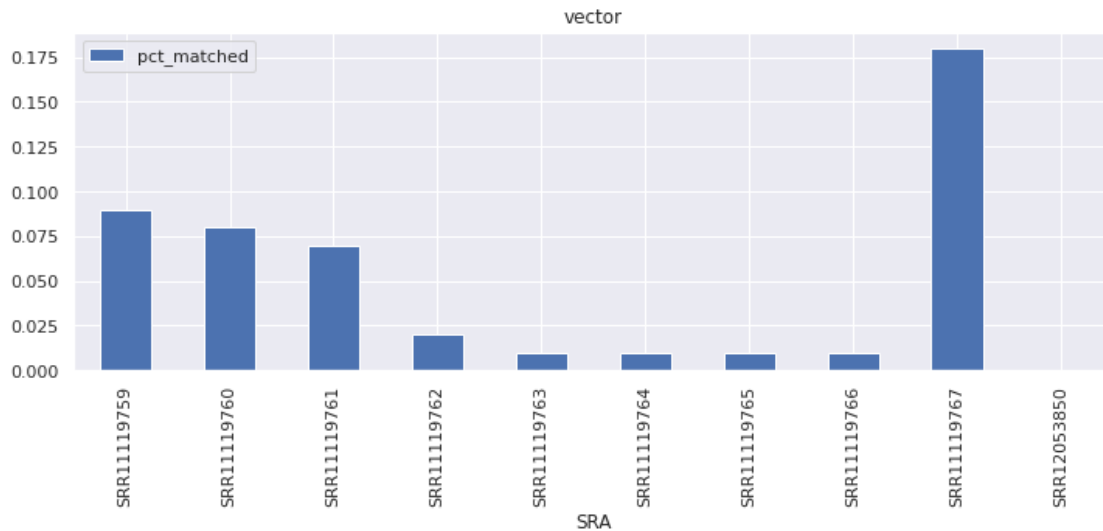
```
[57]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[58]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_vector.png',
↳bbox_inches="tight")
plt.show()
```




```
[59]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_vector_pct.png',
↳bbox_inches="tight")
plt.show()
```



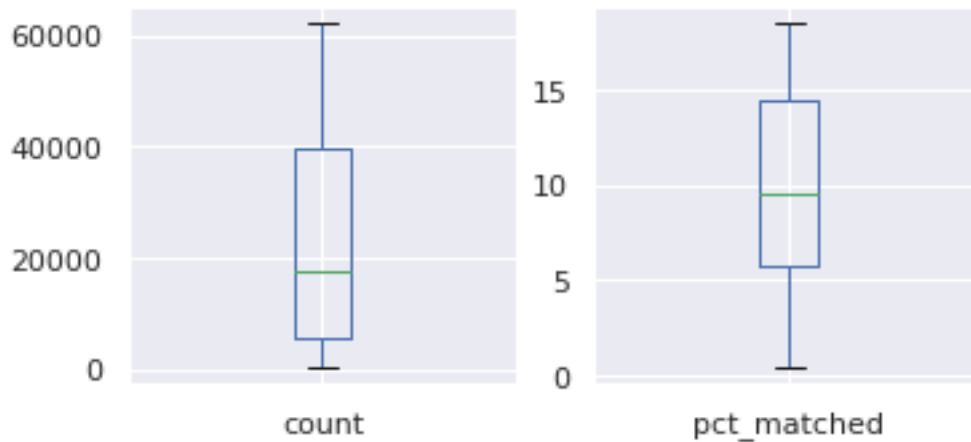
0.0.10 dog

```
[60]: s_name='dog'
```

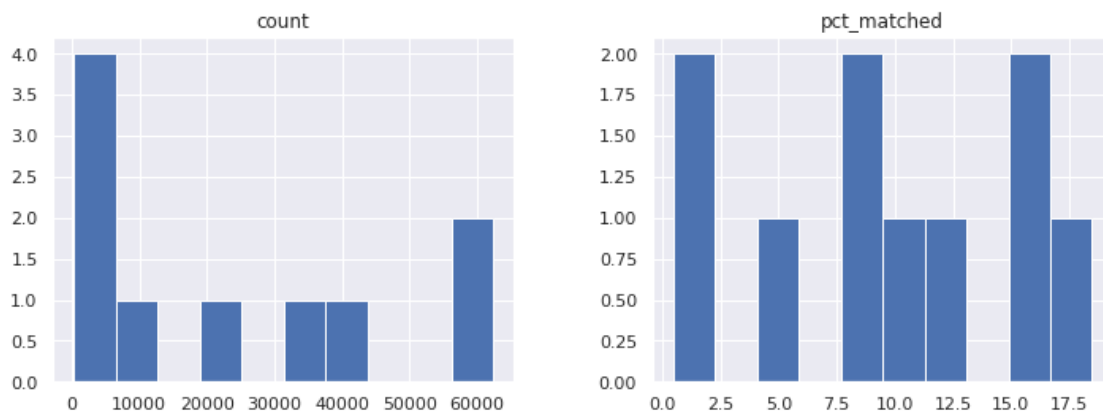
```
[61]: df=df_sra[df_sra['Name'].str.contains(s_name)]
```

```
[62]: df.drop(columns=['Name'],inplace=True)
sns.set(rc={"figure.figsize":(6, 6)})
df_box=df[['count', 'pct_matched']]
df_box.plot(kind='box', subplots=True, layout=(4,4), sharex=False,
↳sharey=False, figsize=(12,12))
```

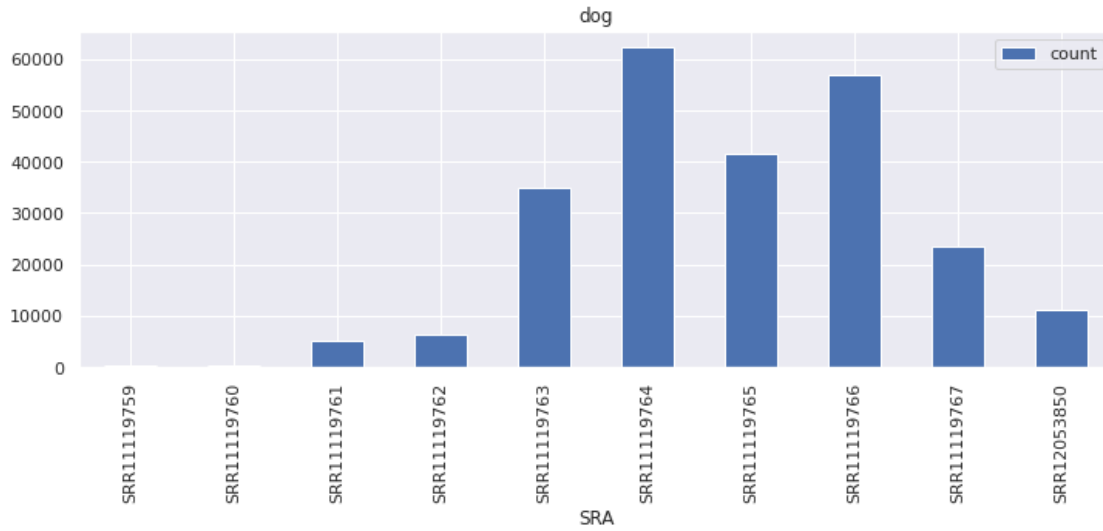
```
[62]: count          AxesSubplot(0.125,0.71587;0.168478x0.16413)
pct_matched        AxesSubplot(0.327174,0.71587;0.168478x0.16413)
dtype: object
```



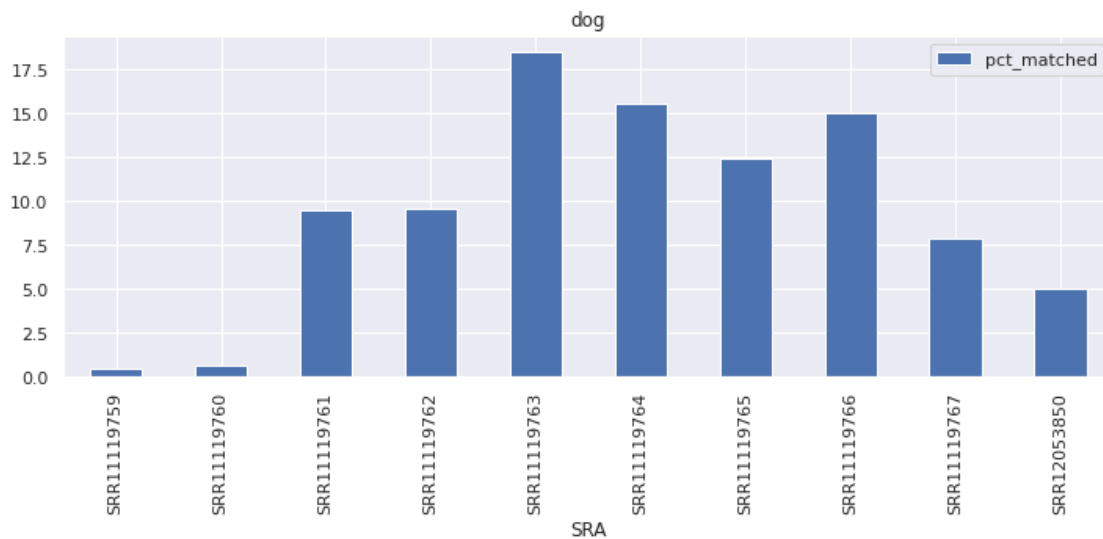
```
[63]: df_box.hist(figsize=(12,4))
plt.show()
```



```
[64]: ax=df.plot(x='SRA', y='count', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_dog.png',
↳bbox_inches="tight")
plt.show()
```



```
[65]: ax=df.plot(x='SRA', y='pct_matched', kind='bar',figsize=(12,4))
ax.set_title(s_name, fontsize=12)
plt.savefig(BASE_PATH+'general_plots/
↳'+f'{PROJECT_CODE}_{dbname}_{kmer}_magicBLAST_contig_dog_pct.png',
↳bbox_inches="tight")
plt.show()
```



0.0.11 Top nt database contigs matches per SRA

```
[66]: def get_sra_dict(sra):
    total_dict={}
    asc_desc={}
    accessions, descriptions, counts, total=get_asc_descr_count(sra)
    for asc,desc,cnt in zip(accessions,descriptions, counts):
        if asc in total_dict:
            total_dict[asc]+=int(cnt)
        else:
            total_dict[asc]=int(cnt)
        if asc not in asc_desc:
            asc_desc[asc]=desc
    return total_dict, asc_desc, total
```

```
[78]: for sra in sra_list:
    total_dict, asc_desc, total= get_sra_dict(sra)
    #f.write(f'{sra}, number of contigs {total}\n')
    print(f'{sra}, number of contigs {total}\n')
    top_n = get_top_n_sp(total_dict, max_num=10)
    for n in top_n:
        print(n)
    print('\n')
```

SRR11119759, number of contigs 33516

Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 16538

Rhodococcus qingshengii CS98 DNA, complete genome: 376

Rhodococcus sp. NJ-530 chromosome, complete genome: 367

Rhodococcus sp. 008, complete genome: 354

Rhodococcus sp. YL-1, complete genome: 332

Rhodococcus sp. djl-6-2 chromosome, complete genome: 320

Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304

Rhodococcus sp. BH4, complete genome: 303

Rhodococcus qingshengii strain 7B chromosome, complete genome: 303

Rhodococcus qingshengii strain IGTS8 chromosome: 295

Rhodococcus sp. P-2 chromosome: 286

Rhodococcus sp. P-2 chromosome: 137

Eukaryotic synthetic construct chromosome 19: 113

Homo sapiens DNA, chromosome 19, nearly complete genome: 104

Eukaryotic synthetic construct chromosome 17: 98

Homo sapiens DNA, chromosome 17, nearly complete genome: 91

Eukaryotic synthetic construct chromosome 19: 85

Eukaryotic synthetic construct chromosome 16: 82

Homo sapiens DNA, chromosome 16, nearly complete genome: 65

Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51

Eukaryotic synthetic construct chromosome 14: 47

Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3470
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2840
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286

SRR11119760, number of contigs 28987

Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 16538
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354

Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 898

Canis lupus familiaris breed Labrador retriever chromosome 14b: 814

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 805

Canis lupus familiaris breed Labrador retriever chromosome 04b: 790

Canis lupus familiaris breed Labrador retriever chromosome 14a: 783

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 6453

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4318

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4174

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 3704

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3644

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 3355

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3205

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2651

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1893

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X2, ncRNA: 1893

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 8121

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 5302

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4951

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781

PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2409

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),

ncRNA: 6825
Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 4530
PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
ncRNA: 4161
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 3668
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 3235
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X2, ncRNA: 2580
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X1, ncRNA: 2525
PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
2424
PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
transcript variant X1, ncRNA: 2376
PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
transcript variant X1, ncRNA: 1867
PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028
PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
ncRNA: 1683
Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1248
PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X2, ncRNA: 771
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367

Rhodococcus sp. 008, complete genome: 354
 Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47

SRR11119761, number of contigs 20964

Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538
 Rhodococcus qingshengii CS98 DNA, complete genome: 376
 Rhodococcus sp. NJ-530 chromosome, complete genome: 367
 Rhodococcus sp. 008, complete genome: 354
 Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523

Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2651

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1893

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X2, ncRNA: 1893

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 8121

Megaerops *niphanae* clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 5302

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201), ncRNA: 4951

Manis tricuspid clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183

Manis tricuspid clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781

PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2409

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825

Megaerops *niphanae* clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161

Manis tricuspid clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668

Manis tricuspid clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2424

PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2376

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1867

PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),

ncRNA: 1683
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1248
PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X2, ncRNA: 771
PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
Eukaryotic synthetic construct chromosome 19: 85
Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551

Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491

SRR11119762, number of contigs 37545

Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538

Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
Eukaryotic synthetic construct chromosome 19: 85
Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151

Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),

ncRNA: 4951
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2409
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2424
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2376
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1867
 PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 1683
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 1273
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 1248
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 1197
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 1055
 Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),

transcript variant X2, ncRNA: 771
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
 Sphingomonas melonis TY, complete genome: 583
 Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
 Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
 Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
 Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
 Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
 Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
 Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
 Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
 Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
 Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
 Rhodococcus qingshengii CS98 DNA, complete genome: 376
 Rhodococcus sp. NJ-530 chromosome, complete genome: 367
 Rhodococcus sp. 008, complete genome: 354
 Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151

Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138

SRR11119763, number of contigs 65302

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538

Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151

Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),

ncRNA: 4951
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2409
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2424
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2376
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1867
 PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 1683
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 1273
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 1248
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 1197
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 1055
 Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),

transcript variant X2, ncRNA: 771
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
 Sphingomonas melonis TY, complete genome: 583
 Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
 Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
 Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
 Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
 Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
 Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
 Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
 Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
 Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
 Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
 Rhodococcus qingshengii CS98 DNA, complete genome: 376
 Rhodococcus sp. NJ-530 chromosome, complete genome: 367
 Rhodococcus sp. 008, complete genome: 354
 Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151

Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
Canis lupus familiaris breed Labrador retriever chromosome 14a: 783

SRR11119764, number of contigs 89894

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 16538
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65

Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3644

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 3355

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3205

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2651

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1893

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X2, ncRNA: 1893

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 8121

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 5302

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4951

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781

PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2409

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2424

PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82

Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),

transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893

SRR11119765, number of contigs 77242

Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 16538
 Rhodococcus qingshengii CS98 DNA, complete genome: 376
 Rhodococcus sp. NJ-530 chromosome, complete genome: 367
 Rhodococcus sp. 008, complete genome: 354
 Rhodococcus sp. YL-1, complete genome: 332
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
 Rhodococcus sp. BH4, complete genome: 303
 Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
 Rhodococcus qingshengii strain IGTS8 chromosome: 295
 Rhodococcus sp. P-2 chromosome: 286
 Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
 Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
 Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
 Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493

Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),

transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3470
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2840
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497

Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X2, ncRNA: 1893
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 8121
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 5302
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201), ncRNA: 4951
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 4183
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2409
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

SRR11119766, number of contigs 86766

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
 Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
 Eukaryotic synthetic construct chromosome 14: 47

Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3470
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2840
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
 Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
 Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
 Eukaryotic synthetic construct chromosome 19: 85
 Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51

Eukaryotic synthetic construct chromosome 14: 47
 Rhodococcus sp. P-2 chromosome: 1772
 Elizabethkingia sp. 2-6 chromosome, complete genome: 597
 Rhodococcus qingshengii CS98 DNA, complete genome: 551
 Rhodococcus sp. YL-1, complete genome: 523
 Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
 Rhodococcus sp. BH4, complete genome: 511
 Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
 Rhodococcus sp. 008, complete genome: 497
 Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
 Rhodococcus sp. NJ-530 chromosome, complete genome: 491
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 375
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
 Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
 Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
 Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
 Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
 Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphanæ clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial

sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3470
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2840
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),

transcript variant X1, ncRNA: 2376
PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
transcript variant X1, ncRNA: 1867

SRR11119767, number of contigs 93945

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 16538

Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
Eukaryotic synthetic construct chromosome 19: 85
Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140

Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
 Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
 Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial

sequence: 3470

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781

PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2409

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA: 2424

PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2376

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1867

PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 1683

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 1273

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 1248

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201), ncRNA: 1197

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 1055

Lutra lutra genome assembly, chromosome: 16: 903

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 771

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 718

Sphingomonas melonis TY, complete genome: 583

Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
Eukaryotic synthetic construct chromosome 19: 85
Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140

Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
 Canis lupus familiaris breed Labrador retriever chromosome 10a: 139
 PREDICTED: Manis javanica angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
 Manis tricusps clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
 Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
 Manis tricusps clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
 Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
 Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
 Manis tricusps clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
 Manis tricusps clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
 Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
 Manis tricusps clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183

Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3470

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2840

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2781

PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2655

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2409

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 2079

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 6825

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 4530

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 4161

Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 3668

Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 3235

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 2580

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 2525

PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA: 2424

PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273), transcript variant X1, ncRNA: 2376

PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430), transcript variant X1, ncRNA: 1867

PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028

PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956), ncRNA: 1683

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 1273

Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial sequence: 1248

PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201), ncRNA: 1197

Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial sequence: 1055

Lutra lutra genome assembly, chromosome: 16: 903

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X2, ncRNA: 771

PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804), transcript variant X1, ncRNA: 718

Sphingomonas melonis TY, complete genome: 583

SRR12053850, number of contigs 79959

Megaerops niphae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial sequence: 16538

Rhodococcus qingshengii CS98 DNA, complete genome: 376

Rhodococcus sp. NJ-530 chromosome, complete genome: 367

Rhodococcus sp. 008, complete genome: 354

Rhodococcus sp. YL-1, complete genome: 332

Rhodococcus sp. djl-6-2 chromosome, complete genome: 320

Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304

Rhodococcus sp. BH4, complete genome: 303

Rhodococcus qingshengii strain 7B chromosome, complete genome: 303

Rhodococcus qingshengii strain IGTS8 chromosome: 295

Rhodococcus sp. P-2 chromosome: 286

Rhodococcus sp. P-2 chromosome: 137

Eukaryotic synthetic construct chromosome 19: 113

Homo sapiens DNA, chromosome 19, nearly complete genome: 104

Eukaryotic synthetic construct chromosome 17: 98

Homo sapiens DNA, chromosome 17, nearly complete genome: 91

Eukaryotic synthetic construct chromosome 19: 85

Eukaryotic synthetic construct chromosome 16: 82

Homo sapiens DNA, chromosome 16, nearly complete genome: 65

Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51

Eukaryotic synthetic construct chromosome 14: 47

Rhodococcus sp. P-2 chromosome: 1772

Elizabethkingia sp. 2-6 chromosome, complete genome: 597

Rhodococcus qingshengii CS98 DNA, complete genome: 551

Rhodococcus sp. YL-1, complete genome: 523

Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516

Rhodococcus sp. BH4, complete genome: 511

Rhodococcus sp. djl-6-2 chromosome, complete genome: 500

Rhodococcus sp. 008, complete genome: 497

Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493

Rhodococcus sp. NJ-530 chromosome, complete genome: 491

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956), ncRNA: 375

Canis lupus familiaris breed Labrador retriever chromosome 04b: 157

Canis lupus familiaris breed Labrador retriever chromosome 06b: 151

Canis lupus familiaris breed Labrador retriever chromosome 06a: 151

Canis lupus familiaris breed Labrador retriever chromosome 04a: 150

Canis lupus familiaris breed Labrador retriever chromosome 10b: 140

Canis lupus familiaris breed Labrador retriever chromosome 08b: 140

Canis lupus familiaris breed Labrador retriever chromosome 08a: 140

Canis lupus familiaris breed Labrador retriever chromosome 10a: 139

PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
ncRNA: 1777

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
transcript variant X2, ncRNA: 1280

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
transcript variant X1, ncRNA: 1256

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
1211

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 948

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 898

Canis lupus familiaris breed Labrador retriever chromosome 14b: 814

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 805

Canis lupus familiaris breed Labrador retriever chromosome 04b: 790

Canis lupus familiaris breed Labrador retriever chromosome 14a: 783

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
ncRNA: 6453

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
ncRNA: 4318

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 4174

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
transcript variant X2, ncRNA: 3704

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 3644

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
transcript variant X1, ncRNA: 3355

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 3205

PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
2651

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
transcript variant X1, ncRNA: 1893

PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
transcript variant X2, ncRNA: 1893

PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
ncRNA: 8121

Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
sequence: 5302

PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
ncRNA: 4951

Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
sequence: 4183

Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
sequence: 3470

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),

transcript variant X2, ncRNA: 2840
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
 Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
 Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: Manis javanica solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: Manis javanica uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376
 PREDICTED: Manis javanica uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: Manis pentadactyla 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: Manis pentadactyla uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
 Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248
 PREDICTED: Manis pentadactyla uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
 Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
 Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: Manis javanica uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
 Sphingomonas melonis TY, complete genome: 583
 Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351
 Canis lupus familiaris breed Labrador retriever chromosome 05b: 344

Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278
Rhodococcus qingshengii CS98 DNA, complete genome: 376
Rhodococcus sp. NJ-530 chromosome, complete genome: 367
Rhodococcus sp. 008, complete genome: 354
Rhodococcus sp. YL-1, complete genome: 332
Rhodococcus sp. djl-6-2 chromosome, complete genome: 320
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 304
Rhodococcus sp. BH4, complete genome: 303
Rhodococcus qingshengii strain 7B chromosome, complete genome: 303
Rhodococcus qingshengii strain IGTS8 chromosome: 295
Rhodococcus sp. P-2 chromosome: 286
Rhodococcus sp. P-2 chromosome: 137
Eukaryotic synthetic construct chromosome 19: 113
Homo sapiens DNA, chromosome 19, nearly complete genome: 104
Eukaryotic synthetic construct chromosome 17: 98
Homo sapiens DNA, chromosome 17, nearly complete genome: 91
Eukaryotic synthetic construct chromosome 19: 85
Eukaryotic synthetic construct chromosome 16: 82
Homo sapiens DNA, chromosome 16, nearly complete genome: 65
Clostridiaceae bacterium 14S0207 chromosome, complete genome: 51
Eukaryotic synthetic construct chromosome 14: 47
Rhodococcus sp. P-2 chromosome: 1772
Elizabethkingia sp. 2-6 chromosome, complete genome: 597
Rhodococcus qingshengii CS98 DNA, complete genome: 551
Rhodococcus sp. YL-1, complete genome: 523
Rhodococcus qingshengii strain RL1 chromosome, complete genome: 516
Rhodococcus sp. BH4, complete genome: 511
Rhodococcus sp. djl-6-2 chromosome, complete genome: 500
Rhodococcus sp. 008, complete genome: 497
Rhodococcus qingshengii strain VKM Ac-2784D chromosome: 493
Rhodococcus sp. NJ-530 chromosome, complete genome: 491
PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
ncRNA: 375
Canis lupus familiaris breed Labrador retriever chromosome 04b: 157
Canis lupus familiaris breed Labrador retriever chromosome 06b: 151
Canis lupus familiaris breed Labrador retriever chromosome 06a: 151
Canis lupus familiaris breed Labrador retriever chromosome 04a: 150
Canis lupus familiaris breed Labrador retriever chromosome 10b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08b: 140
Canis lupus familiaris breed Labrador retriever chromosome 08a: 140
Canis lupus familiaris breed Labrador retriever chromosome 10a: 139

PREDICTED: *Manis javanica* angel homolog 2 (ANGEL2), mRNA: 138
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1777
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 1280
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 1256
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 1211
Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 948
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 898
Canis lupus familiaris breed Labrador retriever chromosome 14b: 814
Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 805
Canis lupus familiaris breed Labrador retriever chromosome 04b: 790
Canis lupus familiaris breed Labrador retriever chromosome 14a: 783
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6453
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4318
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4174
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 3704
Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3644
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 3355
Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3205
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2651
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1893
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X2, ncRNA: 1893
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 8121
Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 5302
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4951
Manis tricuspidis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4183
Manis tricuspidis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3470

PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2840
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2781
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2655
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2409
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 2079
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 6825
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 4530
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 4161
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3668
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 3235
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 2580
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 2525
 PREDICTED: *Manis javanica* solute carrier family 25 member 43 (SLC25A43), mRNA:
 2424
 PREDICTED: *Manis javanica* uncharacterized LOC118973273 (LOC118973273),
 transcript variant X1, ncRNA: 2376
 PREDICTED: *Manis javanica* uncharacterized LOC108385430 (LOC108385430),
 transcript variant X1, ncRNA: 1867
 PREDICTED: *Manis pentadactyla* 28S ribosomal RNA (LOC118908279), rRNA: 2028
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118928956 (LOC118928956),
 ncRNA: 1683
 Megaerops niphanae clone Mnip32b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1273
Manis tricuspis clone Mtri131b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1248
 PREDICTED: *Manis pentadactyla* uncharacterized LOC118913201 (LOC118913201),
 ncRNA: 1197
Manis tricuspis clone Mtri129b LINE-1 retroposon ORF II pseudogene, partial
 sequence: 1055
Lutra lutra genome assembly, chromosome: 16: 903
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X2, ncRNA: 771
 PREDICTED: *Manis javanica* uncharacterized LOC108383804 (LOC108383804),
 transcript variant X1, ncRNA: 718
Sphingomonas melonis TY, complete genome: 583
Felis catus Senzu DNA, chromosome: E1, American Shorthair breed: 351

```

Canis lupus familiaris breed Labrador retriever chromosome 05b: 344
Canis lupus familiaris breed Labrador retriever chromosome 05a: 328
Felis catus Senzu DNA, chromosome: D4, American Shorthair breed: 305
Canis lupus familiaris breed Labrador retriever chromosome 02a: 301
Canis lupus familiaris breed Labrador retriever chromosome 02b: 299
Canis lupus familiaris breed Labrador retriever chromosome 20b: 292
Canis lupus familiaris breed Labrador retriever chromosome 20a: 287
Felis catus Senzu DNA, chromosome: E2, American Shorthair breed: 286
Felis catus Senzu DNA, chromosome: D2, American Shorthair breed: 278

```

```
[68]: ### Get specific contigs matching a species/name
```

```
[69]: def write_contigs(sra, match_names):
    accessions, descriptions, counts, total=get_asc_descr_count(sra)
    asc_matches=[]
    for m in match_names:
        for a,d in zip(accessions, descriptions):
            if m.lower() in d.lower():
                asc_matches.append(a)
    gi_matches=[]
    for a in asc_matches:
        idx=ACCESSIONS.index(a)
        gi=GIS[idx]
        gi_matches.append(gi)
    contigs=[]
    path = BASE_PATH+sra+'/magic_blast/'
    subset_f = open(path+f'{sra}_{match_names[0].replace(" ", "_")}_subset_{f_contigs_file_tail}', 'w')
    with open(path+f'{sra}_{f_contigs_file_tail}', 'r') as f:
        lines = [line for line in f]
        for line in lines:
            for gi in gi_matches:
                if gi in line:
                    parts=line.split('\t')
                    idx=GIS.index(gi)
                    asc=ACCESSIONS[idx]
                    t=TITLES[idx]
                    ps=parts[:2]
                    pe=parts[3:]
                    ps.append(asc+' '+t.rstrip('\n'))
                    parts=ps+pe
                    p='\t'.join(parts)
                    subset_f.write(p)
    subset_f.close()
```

```
[ ]:
```

```
[70]: #set_accessions()
      #assert ACCESSIONS is not None
```

```
[71]: #match_names=['virus']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[72]: #match_names=['vector']
      #for sra in sra_short_list:
      #    write_contigs(sra, match_names)
```

```
[73]: #match_names=['plasmid']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[74]: #match_names=['mustela']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[75]: #match_names=['chlorocebus']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[76]: #match_names=['dog']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[77]: #match_names=['canis lupus']
      #for sra in sra_list:
      #    write_contigs(sra, match_names)
```

```
[ ]:
```